Proceedings of the Sixth International Conference on
Railway Technology: Research, Development and Maintenance
Edited by: J. Pombo
Civil-Comp Conferences, Volume 7, Paper 3.11
Civil-Comp Press, Edinburgh, United Kingdom, 2024
ISSN: 2753-3239, doi: 10.4203/ccc.7.3.11

Application of Kriging Model with Sequential Infill Criterion on Multi-Objective Optimization of Nose Shape for High-Speed Train

Z. Dai[1,2], T. Li[1], S. Krajnovic[2] and W. Zhang[1]

[1]State Key Laboratory of Rail Transit Vehicle System, Southwest Jiaotong University, Chengdu, China
[2]Department of Mechanics and Maritime Sciences, Chalmers University of Technology, Gothenburg, Sweden

Abstract

This study proposes a sequential infill criterion (SIC) appropriate for the Kriging surrogate to address this issue. Multi-objective functions are employed to test the feasibility of constructing a surrogate model based on SIC, and the SIC surrogate model then performs multi-objective aerodynamic optimizations on the high-speed train. The findings indicate that the improvement infill criterion (EIC) that fuses the gradient information (PGEIC) surrogate model achieves the lowest generational distance (GD) and prediction error. The performance of EIC for global search, EIC for Pareto front search (PEIC), and infill criterion for Pareto front search using only gradient information (PGIC) is poor. The final PGEIC-SIC surrogate model of train aerodynamics has less than 1% prediction error for the three optimization objectives. The optimal solution reduces the aerodynamic drag force of the head car and the aerodynamic drag and lift force of the tail car by 4.15%, 3.21%, and 3.56%, respectively, compared with the original model. Furthermore, the nose and cab window heights of the optimal model have been reduced, and the lower contour line is concave. Correspondingly, the streamlined shape appears more rounded and slender.

Keywords: surrogate model, sequential infill criterion, gradient information, high-speed train, aerodynamic multi-optimization, shape optimization.

1 Introduction

Aerodynamic multi-objective optimization of high-speed trains is important for improving train operation safety and riding comfort. However, the geometric model of the high-speed train has a large scale and complex structure. The computing cost of numerical simulation is huge, lengthening the multi-objective optimization cycle and making it harder to find the global optimal solution. As a result, optimization using surrogate models emerged as a critical component for accelerating the optimization design and enhancing optimization efficiency. The prediction accuracy of the surrogate model seriously affects the optimization results when aerodynamic design is performed. Appropriate sample infill criteria can improve the model accuracy and facilitate the convergence of multi-objective optimization [1]. The traditional surrogate model method without infill criteria has been widely used in high-speed train aerodynamic optimization [2-5]. The progress of numerical technology has concurrently fostered an inclination towards addressing multiple variables and objectives, leading to a notable escalation in the requisite number of samples essential for the establishment of high-precision surrogate models. Zhang et al. [6] proposed a method of infilling special samples in the Pareto front, selecting the optimal point of every objective for verification. If the error criterion is not met, it will be added to the training set to refine the model. Xu et al. [7] established a cross-validation (CV) Kriging surrogate model. Should the prediction error of the validation sample fail to meet the predetermined criteria, it is incorporated into the training set. Subsequently, a succeeding-generation model is formulated to continue the optimization process. Yao et al. [8] and other researchers [9] constructed standard Kriging or CV-Kriging models and defined error standards. The Pareto front was then obtained using optimization algorithms, and validation samples were selected. However, the papers don't contain details on how to select samples from the optimized solution set for verification. The approach of selecting points in the Pareto front is widely used and has successful outcomes. It can make the model converge to the optimal solution faster and improve the optimization efficiency [9-11]. In this study, the sequential infill criteria (SIC) is introduced, a novel approach that amalgamates the gradient information derived from the surrogate model with the expected improvement infill criterion (EIC). During the initial stage of the SIC, the EIC is employed to enhance the global accuracy of the surrogate model. Then, the modeling process transitions to the second stage upon convergence of prediction accuracy. Subsequently, infill criteria grounded in the Pareto front are applied to enhance prediction accuracy near the optimal solution until the predetermined standard is met, thereby identifying the optimal solution. The application of the SIC serves to enhance the accuracy of the surrogate model with minimal sample utilization, consequently elevating the efficiency of aerodynamic multi-objective optimization. This methodology is indispensable for addressing multi-objective optimization problems characterized by extensive numerical computations.

2 Methods

2.1 Sequential infill criterion (SIC)

The construction process of the SIC surrogate model, as illustrated in Figure 1, unfolds as follows:

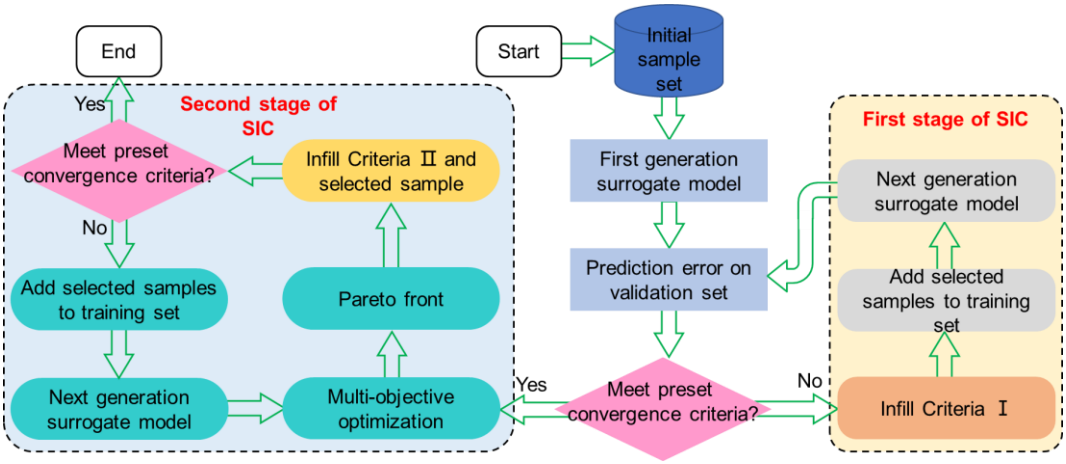


Figure 1: Construction process of the SIC surrogate model.

Step 1: The initial sample set is derived through the optimal Latin hypercube sampling (OLHS) method. Meanwhile, the sample expansion strategy in the literature [12] is used to obtain the training and validation sets, facilitating the construction of the initial surrogate model.

Step 2: The validation set is employed to assess the prediction accuracy of the surrogate model. The average prediction error of the validation set serves as the evaluation index for this purpose. If the prediction error fulfills the predefined requirements, the global accuracy is regarded to meet the requirements, and the second stage is entered.

Step 3: If the prediction error does not reach the predefined requirements, the sample infill criterion I EIC is adopted to determine the samples to be added. Add the sample to the training set, carry out training of the next-generation surrogate model, and return to *Step* 2.

Step 4: During the second stage of the SIC, the NSGA-II is utilized for calculating the Pareto front, and sample infill criterion II is employed to explore the Pareto front, determining the samples to be added.

Step 5: The chosen samples are employed to validate the prediction accuracy of the surrogate model. Should the prediction error meet the predefined requirements, it is concluded that the surrogate model and the multi-objective optimization are sufficiently accurate, leading to the termination of the optimization procedure.

Step 6: If the prediction error does not satisfy the predefined criteria, the selected sample from the Pareto front is added to the training set. Then, the training of the next-generation surrogate model is carried out and returned to *Step* 4.

In Step 1, the sample expansion strategy identifies samples from the initial sample set using the maximum and minimum distance criterion combined with the optimization technique. The obtained validation set contains no samples that overlap with the training set. Furthermore, it can ensure the maximum sample filling degree, enabling the validation set to test the surrogate model's global error.

The convergence criterion S_{thr}^2 is the variance of the prediction error of the last three generations of the validation set. When S_{thr}^2 <0.01, the global accuracy is regarded to have converged and moves to the second stage of SIC.

$$S_{thr}^2 = \frac{1}{3}\sum_{i=n-2}^{n}\left(e_i - \overline{e}\right)^2 \tag{1}$$

where n is the generation number, e_i is the prediction error of the i^{th} generation surrogate model on the validation set. $\overline{e}$ is the average prediction error of the last three generations of the validation set.

In the second stage of SIC, the construction efficiencies of surrogate models for four infill criteria are compared: the EIC for global search, the EIC for Pareto front search (PEIC), the infill criterion for Pareto front search using only gradient information (PGIC), and an infill criterion based on EIC that fuses gradient information (PGEIC). Particularly, the PGEIC finds the sample with the largest gradient g in the Pareto front and adds it to the training set to generate the next-generation surrogate model. PGEIC integrates the EIC and the gradient information of the sample, as illustrated in Equation (3). The sample characterized by the maximum I_g is incorporated into the training set to generate the next-generation surrogate model.

$$g = \sum_{i=1}^{n}\sqrt{\sum_{j=1}^{k}\left(\frac{\partial y_i}{\partial x_j}\right)^2} \tag{2}$$

$$I_g = \frac{g}{g_{\max}} I \tag{3}$$

where n is the objective dimension, k is the variable dimension

2.2 Setting of numerical simulations

Figure 2 shows the numerical simulation domain, with the height and width of 14H and 28H, respectively. The x distances between the nose tip and the boundaries of the domain are 14H and 28H, respectively, which satisfy the standards of Requirements and test procedures for aerodynamics on open track (EN 14067-4). The rail and subgrade are built to simulate the real ground configuration. The origin of the coordinates is situated at the bottom of the middle car, aligned with the longitudinal center section of the car body. Set up a refinement box to refine the grid surrounding the train to solve the flow field more accurately.

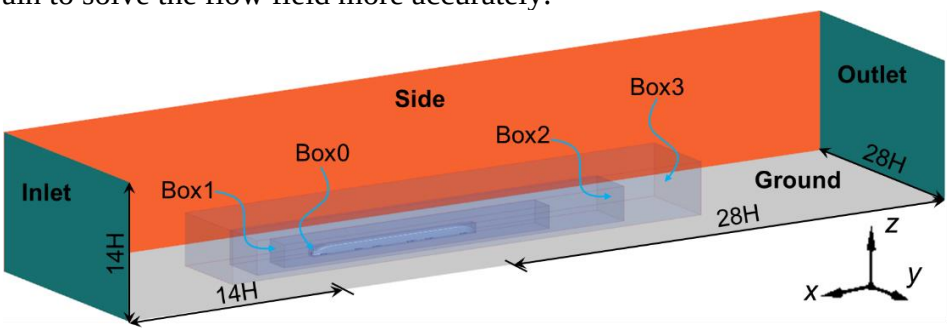


Figure 2: Computational domain.

The velocity-inlet boundary is specified, with velocity $\boldsymbol{u}$ = (97.22, 0, 0) m/s. The airflow of the surroundings becomes a compressible, viscous, unsteady turbulent flow during high-speed train operation. In the present work, the flow is computed using the Reynolds average method, along with the k-ω SST (shear stress transport) two-equation turbulence model [13-14]. This is done to reduce the computational cost of the optimization process. Outlet boundary is designated as the pressure-outlet, with the pressure set to 0. The top and sides of the computational domain are configured as symmetry.

2.3 Parametric design of train nose shape

A full-scale three-car marshaling model, comprising the head, middle, and tail cars, is employed for the multi-objective optimization of the high-speed train, as depicted in Figure 3. The objectives in the multi-objective aerodynamic optimization of high-speed trains include the aerodynamic lift force of the tail car (LT), the aerodynamic drag force of the tail car (DT), and the aerodynamic drag force of the head car (DH). The DH, DT, and LT of the original model are 6.98 kN, 4.68 kN, and 3.65 kN, respectively.

Five features are selected to parameterize the high-speed train, namely control line of nose height C_1, gear height C_2, cab window height C_3, middle contour line C_4, and lower contour line C_5, corresponding to design variables v_1、v_2、v_3、v_4 and v_5. v_1 represents the deformation ratio of the original curve, v_2、v_3、v_4, and v_5 represents the maximum distance of the curve deformation (mm). Figure 3 depicts the design variables and change intervals. The curves are all B-spline curves derived from numerous design points, and the surfaces are B-spline surfaces formed from spline curves.

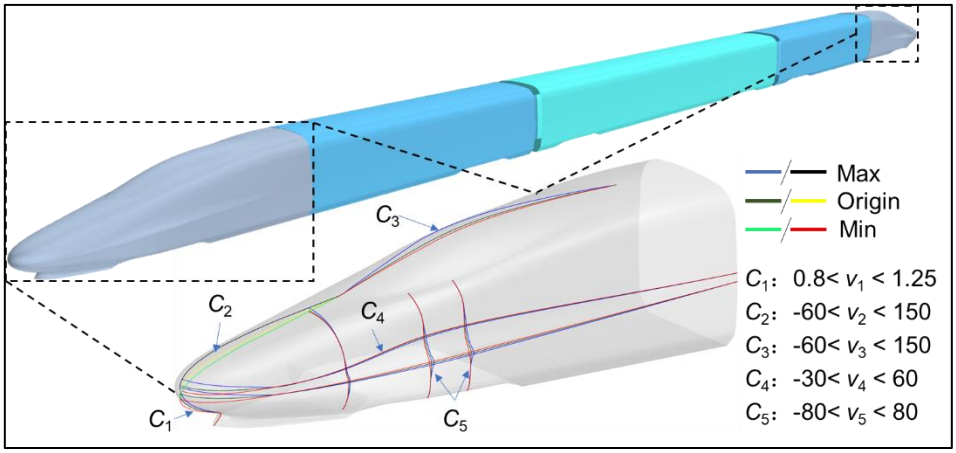


Figure 3: Design variables and variation intervals.

3 Results and discuss

3.1 ZDT1 Test functions

The surrogate model construction efficiency of the four criteria is compared using the multi-objective test functions Zitzler-Deb-Thiele 1 (ZDT1). This comparison aims to establish the universal applicability of the PGEIC-SIC surrogate model.

The f_1 and f_2 of the ZDT1 [15] are:

$$f_1(x) = x_1 \tag{4}$$

$$f_2(x) = g(x)h\left(f_1(x), g(x)\right) \tag{5}$$

where $g(x) = 1 + \frac{9}{29}\sum_{i=2}^{30} x_1$, $h\left(f_1(x), g(x)\right) = 1 - \sqrt{\frac{f_1(x)}{g(x)}}$, $0 \le x_i \le 1, 1 \le i \le 30$.

The initial training sets for ZDT1 consist of 15 and 30 samples, respectively. The convergence requirements of the SIC first stage are satisfied after adding 11 and 18 samples, respectively, using EIC, and progress to the second stage of SIC. Figure 4 depicts the GD obtained in the SIC second step using various sample infill criteria. The GD obtained from the PGEIC surrogate model has the smallest value for two functions, indicating that the error between it and the real Pareto front is the smallest, followed by the PEIC. The EIC method yields the highest GD among the four methods. This is primarily due to the high variable dimension of the ZDT1 function, which makes it difficult to get close to the optimal solution when performing a global search. Therefore, the created surrogate model has a larger prediction error near the real Pareto front or even far away from it. The analysis of the second stage infill criteria of SIC reveals that PGEIC has the greatest efficiency in constructing a surrogate model, whereas EIC and PGIC have lower efficiencies. Consequently, PGEIC will be used to create the surrogate model, the PGEIC-SIC surrogate model, on which the aerodynamic multi-objective optimization research on high-speed trains will be conducted.

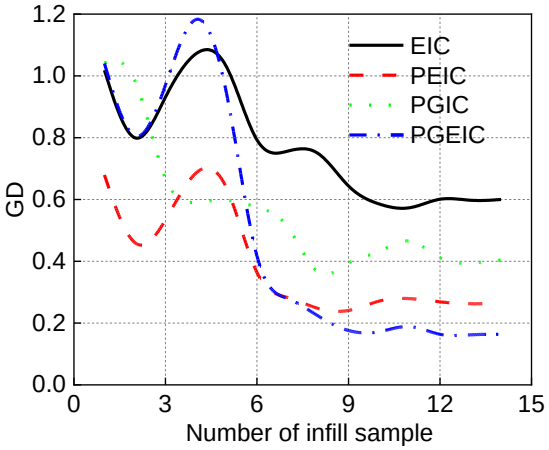
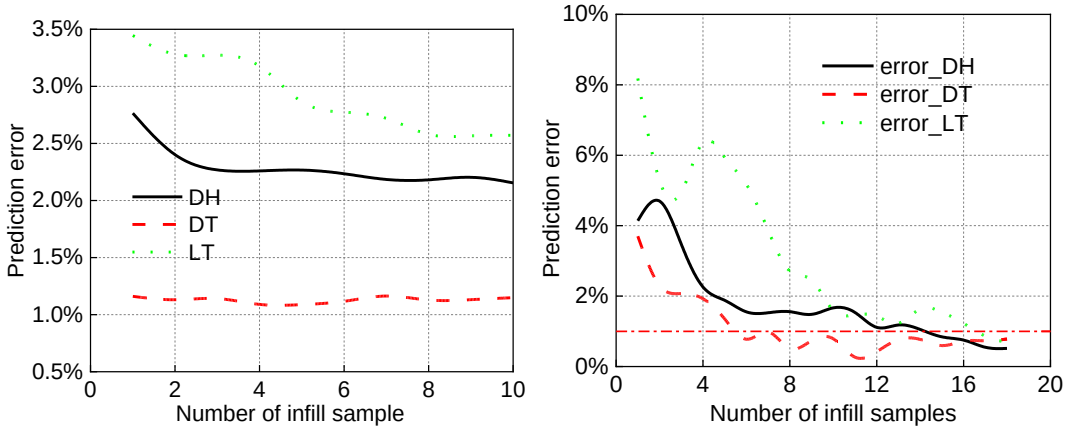


Figure 4: GD coverage history of four surrogate models in the SIC second stage.

3.2. Construction of SIC surrogate model and multi-optimization

The OLHS method is employed to generate 40 samples and establish an initial sample set. Ten samples are acquired through the sample expansion strategy, constituting a validation set, and the remaining 30 samples are designated as the training set. The first-generation surrogate model is constructed using the initial training set, with average prediction errors of 2.8%, 1.2%, and 3.5% for DH, DT, and LT, respectively. The first stage of the SIC surrogate model is executed following the procedural steps delineated in Figure 1. After adding the first sample, the errors of the new generation model for the three objectives of the validation set are 2.4%, 1.1%, and 3.2%, respectively. The global prediction error convergence process is depicted in Figure

5(a). The relative error is adopted as the error evaluation criterion, and the variance of the data is small. Therefore, the predetermined criterion is reduced to 10^{-6}. The values of the three objective functions are 8.4×10^{-8}, 2.4×10^{-8}, and 3.7×10^{-8} when 10 samples are infilled and the first stage finishes.



(a) Globel prediction error (b) Prediction error for validation samples

Figure 5: Prediction error of the PGEIC-SIC surrogate model.

The aerodynamic multi-objective optimizations of high-speed trains are an actual engineering problem, which differs from the optimization of multi-objective test functions. It is impossible to determine the convergence of optimization using GD because there is no recognized Pareto front. The ultimate goal of multi-optimization for actual engineering issues is to obtain an optimized solution with better performance. Therefore, the prediction error of the sample in the Pareto front is chosen as the criterion for determining if optimization is finished. The termination criterion for multi-objective optimization in this study is that DH, DT, and LT prediction errors are below 1% simultaneously. The NSGA-II algorithm is engaged to figure out the Pareto front, and then the first sample to be added in the second stage of SIC is determined using PGEIC. The relative errors of the surrogate model for the DH, DT, and LT are 4.14%, 3.70%, and 8.18% when the model is updated. Since the errors of the surrogate model fail to match predefined standards, the validation sample is incorporated into the training set for generating the next-generation model. This iterative process is repeated until the prediction errors of the surrogate model for DH, DT, and LT are less than 1%. The optimization convergence process of the PGEIC-SIC surrogate model is shown in Figure 5(b). The prediction errors of objectives are less than 1% as the second-stage surrogate model is updated to the 18th generation. It is apparent that the convergence of the prediction error for DL is faster, whereas DH and LT converge a little bit slower.

The validation sample selected from the Pareto front is chosen as the optimal solution of the multi-objective aerodynamic optimization. The parameters v_1, v_2, v_3, v_4, and v_5 of the optimal solution are 0.83, 131.96, -33.80, 55.17, and -69.07, corresponding to the 1.00, 0, 0, 0, and 0 of the original model. The optimal model exhibits reduced heights for the nose and cab window, coupled with an increased height in the region between the two sections, as depicted in Figure 6. This alteration contributes to a smoother head shape. The middle contour line is convex, while the

bottom contour line is concave, as indicated by the positive value of v_4 and the negative value of v_5. Consequently, the cross-section of the head is more rounded, and the longitudinal shape is more slender.

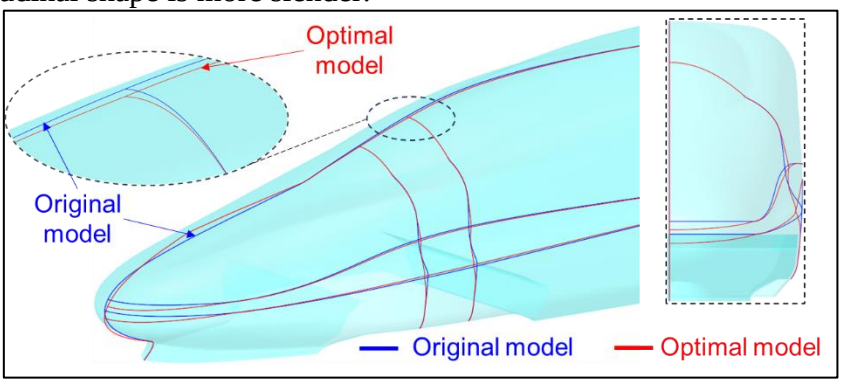
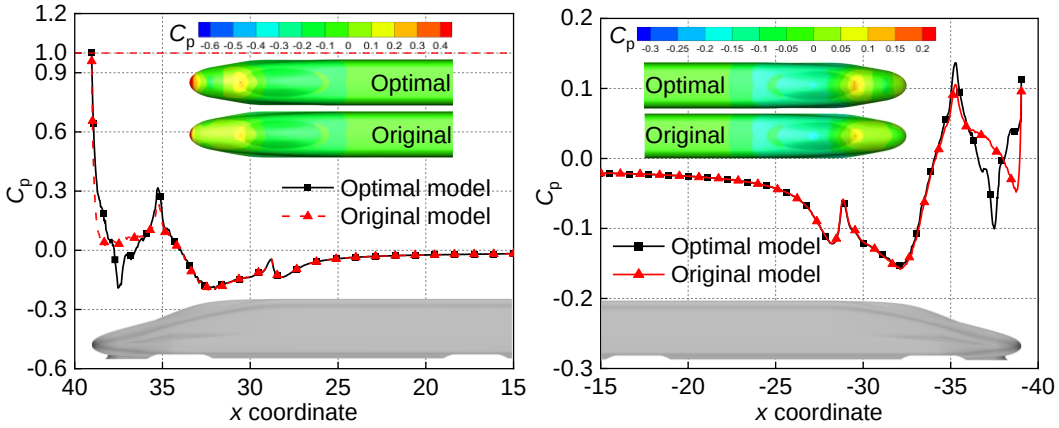


Figure 6: Comparison between the geometry models of the original and optimal.

The DH, DT, and LT of the optimal model are 6.69 kN, 4.53 kN, and 3.52 kN, respectively, which are reduced by 4.15%, 3.21%, and 3.56% from the original model. Each objective of the optimal model improves performance by roughly 3% to 4%, with the DH improving the most. To further analyze the changes in aerodynamic performance between the optimal and original train models, as shown in Figure 7, the surface pressure coefficient of the carriages and the pressure coefficient of the longitudinal center plane of the train are extracted. The surface pressure difference between the optimal and original models is mostly centered in the region from the nose tip to the cab window of the head car, where the geometric model variations are greatest. It is clear that the pressure distribution of the two models from the cab window to downstream essentially overlaps, indicating that there is almost no difference in surface pressure.



(a) Head car (b) Tail car

Figure 7: Pressure coefficient of the head and tail car.

The surface pressure suggests that the zone from the nose tip of the original model to the cab window is under positive pressure, which is confirmed by the mid-section pressure data. However, the optimal model generates a part of negative pressure behind the nose tip. The traction force produced by this negative pressure reduces the aerodynamic resistance of the optimal model. which is the primary reason why the

optimal model has a lower aerodynamic resistance. Moreover, the nose stagnation pressure coefficients of the optimal and original models are identical to 1, confirming the accuracy of the numerical simulation. Furthermore, surface pressure between the body and the upstream part of the cab window differs little between the two models in the tail car. The fundamental distinction is still concentrated in the region between the cab window and the nose tip, and the positive pressure amplitudes at the cab window and the top of the nose tip for the optimal model are larger. Meanwhile, the negative pressure of the nose tip upstream is also larger. The positive pressure will weaken the aerodynamic resistance and aerodynamic lift of the tail car, while negative pressure has the opposite effect. Under the combined influence of the pressure difference in the streamlined region of the tail car, the aerodynamic resistance and aerodynamic force of the optimal model are slightly less than the original model

4 Conclusions and Contributions

A sequential infill criterion (SIC) suitable for the Kriging surrogate model is proposed in this study. Multi-objective optimization functions are used to test the SIC, which verified the feasibility of establishing an SIC surrogate model. Finally, multi-objective aerodynamic optimization on the high-speed train using the SIC surrogate model is performed, and sensitivity analysis is conducted. The following are the primary conclusions:

ZDT1 test results suggest that utilizing EIC in the first stage of the SIC surrogate model can improve global prediction. The Kriging surrogate model generated by the EIC integrating gradient information (PGEIC) has the highest prediction accuracy in the second stage. The PGEIC-SIC surrogate model has the smallest GD index and the highest optimization efficiency. The performance of EIC for Pareto front search (PEIC), EIC for global search decreases gradually, and the infill criterion for Pareto front search using only gradient information (PGIC) performs the worst.

An aerodynamic PGEIC-SIC Kriging surrogate model is constructed using 5 key parameters of the nose shape and three objectives of DH, DT, and LT. The average prediction errors of the final surrogate model of the validation samples are less than 1%. The optimal solution's DH, DT, and LT are 4.15%, 3.21%, and 3.56% lower than the original model. The nose and cab window heights of the optimal model are reduced. Meanwhile, the middle contour line is convex, while the bottom contour line is concave. The cross-section of the head is more rounded, and the longitudinal shape is more slender.

Acknowledgements

We are grateful for support from SNIC (Swedish National Infrastructure for Computing) at the National Supercomputer Center (NSC) at LiU, acknowledging the financial support from China Scholarship Council (Grant No. 202307000087).

References

[1] Z. H. Han, “Kriging surrogate model and its application to design optimization: A review of recent progress,” Acta Aeronautica et Astronautica Sinica, 37(11): 3197-3225, 2016.

[2] Z. Sun, J. Song, and Y. An, “Optimization of the head shape of the CRH3 high-speed train,” Science China Technological Sciences, 53, 3356-3364, 2010.

[3] S. B. Yao, D. L. Guo, Z. X. Sun, G. W. Yang, and D. W. Chen, “Optimization design for aerodynamic elements of high-speed trains,” Computers and Fluids, 95, 56-73, 2014.

[4] N. Zhang, P. Wang, H. Dong, and T. Li, “Shape optimization for blended-wing–body underwater glider using an advanced multi-surrogate-based high-dimensional model representation method,” Engineering Optimization, 52(12), 2080-2099, 2020.

[5] J. Muñoz-Paniagua, J. García, and A. Crespo, “Genetically aerodynamic optimization of the nose shape of a high-speed train entering a tunnel,” Journal of Wind Engineering and Industrial Aerodynamics, 130, 48-61, 2014.

[6] L. Zhang, J. Y. Zhang, T. Li, and W. H. Zhang, “Optimization on the crosswind stability of trains using neural network surrogate model,” Chinese Journal of Mechanical Engineering, 34(4): 208-224, 2021.

[7] G. Xu, X. Liang, S. Yao, D. Chen, and Z. Li, “Multi-objective aerodynamic optimization of the streamlined shape of high-speed trains based on the Kriging model,” PloS one, 12(1), e0170803, 2017.

[8] S. B. Yao, D. L. Guo, Z. X. Sun, and G. W. Yang, “A modified multi-objective sorting particle swarm optimization and its application to the design of the nose shape of a high-speed train,” Engineering Applications of Computational Fluid Mechanics, 9(1), 513-527, 2015.

[9] S. B. Yao, D. L. Guo, Z. X. Sun, G. W. Yang, and D. W. Chen, “Optimization design for aerodynamic elements of high-speed trains,” Computers and Fluids, 95, 56-73, 2014.

[10] L. Zhang, J. Zhang, T. Li, and Y. Zhang, “A multi-objective aerodynamic optimization design of a high-speed train head under crosswinds,” Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, 232(3), 895-912, 2018.

[11] Z. He, T. Liu, and H. Liu, “Improved particle swarm optimization algorithms for aerodynamic shape optimization of high-speed train,” Advances in Engineering Software, 173, 103242, 2022.

[12] Z. Y. Dai, T. Li, Z. R. Xiang, W. H. Zhang, and J. Y. Zhang, “Aerodynamic multi-objective optimization on train nose shape using feedforward neural network and sample expansion strategy,” Engineering Applications of Computational Fluid Mechanics, 17(1), 2226187, 2023.

[13] T. Li, Z. Y. Dai, M. Yu, and W. H. Zhang, “Numerical investigation on the aerodynamic resistances of double-unit trains with different gap lengths,” Engineering Applications of Computational Fluid Mechanics, 15(1), 549-560, 2021.

[14] J. Zhang, A. Adamu, F. Gidado, M. Tang, O. Ozer, and X. Chen, “Flow control for aerodynamic drag reduction of a high-speed train with diversion slots on bogie regions,” Physics of Fluids, 35, 115111, 2023.
[15] E. Zitzler, L. Thiele, M. Laumanns, C. M. Fonseca, and V. G. Fonseca, “Performance assessment of multi-objective optimizers: An analysis and review,” IEEE Transactions on Evolutionary Computation, 7(2), 117-132, 2003.